

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Thomas, Terry L.

(ii) TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A  
DELTA 6-DESATURASE

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Scully, Scott, Murphy & Presser
- (B) STREET: 400 Garden City Plaza
- (C) CITY: Garden City
- (D) STATE: New York
- (E) COUNTRY: United States
- (F) ZIP: 11530

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Presser, Leopold
- (B) REGISTRATION NUMBER: 19,827
- (C) REFERENCE/DOCKET NUMBER: 8383ZYXWVU

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (516) 742-4343
- (B) TELEFAX: (516) 742-4366
- (C) TELEX: 230 901 SANS UR

FOI 2025 9562007

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3588 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2002..3081

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCTAGCCACC AGTGACGATG CCTTGAATTT GGCCATTCTG ACCCAGGCCC GTATTCTGAA	60
TCCCCGCATT CGCATTGTTA ATCGTTTGTT CAACCATGCC CTGGGTAAAC GTTTAGACAC	120
CACCTTGCCA GACCACGTTA GTTTGAGTGT TTCCGCCCTG GCGGCCCCGA TTTTTCCTT	180
TGCGGCTTTG GGCAATCAGG CGATCGGGCA ATTGCGTTTG TTTGACCAGA CTTGGCCCAT	240
TCAGGAAATT GTCATTCACC AAGACCATCC CTGGCTCAAT TTACCCCTGG CGGATTTATG	300
GGATGATCCG AGCCGAATGT TGATCTATTA CCTACCGGCC CACAGTGAAA CGGATTTAGT	360
AGGCGCAGTG GTGAATAATT TAACGTTGCA ATCTGGGGAC CATTTAATAG TGGGACAAAA	420
ACCCCAACCC AAGACCAAAC GGCGATCGCC TTGGCGCAAA TTTTCCAAAC TGATTACCAA	480
CCTGCGGGAG TATCAGCGGT ATGTCCAACA GGTGATATGG GTGGTGTGTG TTTTATTGTT	540
GATGATTTTT CTGGCCACCT TCATCTACGT TTCCATTGAT CAACATATTG CCCCAGTGGA	600
CGCGTTGTAT TTTTCCGTGG GCATGATTAC CGGGGCCGGT GGCAAGGAAG AGGTGGCCGA	660
AAAGTCCCCC GATATCATCA AAGTATTCAC AGTGGTGATG ATGATCGCCG GGGCGGGGGT	720
GATTGGTATT TGTTATGCCC TACTGAATGA TTTCATCCTT GGCAGTCGCT TTAGTCAGTT	780
TTTGGATGCG GCCAAGTTAC CCGATCGCCA TCACATCATC ATTTGTGGGC TGGGGGGAGT	840
GAGCATGGCC ATTATTGAAG AGTTAATTCA CCAGGGCCAT GAAATTGTGG TAATCGAAAA	900

GGATACAGAT AATCGTTTCT TGCATACGGC CCGCTCCCTG GGGGTGCCCG TAATTGTGGA 960  
GGATGCCCCG CTAGAAAGAA CGTTGGCCTG CGCCAATATC AACCGAGCCG AAGCCATTGT 1020  
GGTGGCCACC AGCGACGACA CCGTTAACTT GGAAATTGGC CTAAGTCCA AGGCGATCGC 1080  
CCCTAGCCTG CCAGTGGTGT TGC GTTGCCA GGATGCCAG TTTAGCCTGT CCCTGCAGGA 1140  
AGTATTTGAA TTTGAAACGG TGCTTTGTCC GCGGGAATTG GCCACCTATT CCTTTGCGGC 1200  
GGCGGCCCTG GGGGGCAAAA TTTTGGGCAA CGGCATGACC GATGATTTGC TGTGGGTAGC 1260  
CCTAGCCACC TTAATCACTC CTAACCATCC CTTTGCCGAC CAATTGGTTA AAATTGCAGC 1320  
CCAAAAGTCT GATTTCGTTC CCCTCTATCT AGAACGGGGT GGCAAAACCA TCCATAGCTG 1380  
GGAATTATTG GGTACCCATC TCGACTCTGG AGACGTGTTG TATTTAACCA TGCCCGCCAC 1440  
TGCCCTAGAG CAACTTTGGC GATCGCCCCG TGCCACTGCT GATCCTCTGG ACTCTTTTTT 1500  
GGTTTAGCAT GGGGGGATGG AACTCTTGAC TCGGCCCAAT GGTGATCAAG AAAGAACGCT 1560  
TTGTCTATGT TTAGTATTTT TAAGTTAACC AACAGCAGAG GATAACTTCC AAAAGAAATT 1620  
AAGCTCAAAA AGTAGCAAAA TAAGTTTAAT TCATAACTGA GTTTTACTGC TAAACAGCGG 1680  
TGCAAAAAAG TCAGATAAAA TAAAAGCTTC ACTTCGGTTT TATATTGTGA CCATGGTTCC 1740  
CAGGCATCTG CTCTAGGGAG TTTTCCGCT GCCTTTAGAG AGTATTTTCT CCAAGTCGGC 1800  
TAACTCCCCC ATTTT TAGGC AAAATCATAT ACAGACTATC CCAATATTGC CAGAGCTTTG 1860  
ATGACTCACT GTAGAAGGCA GACTAAATTT CTAGCAATGG ACTCCCAGTT GGAATAAATT 1920  
TTTAGTCTCC CCCGGCGCTG GAGTTTTTTT GTAGTTAATG GCGGTATAAT GTGAAAGTTT 1980  
TTTATCTATT TAAATTTATA A ATG CTA ACA GCG GAA AGA ATT AAA TTT ACC 2031  
Met Leu Thr Ala Glu Arg Ile Lys Phe Thr  
1 5 10  
CAG AAA CGG GGG TTT CGT CGG GTA CTA AAC CAA CGG GTG GAT GCC TAC 2079  
Gln Lys Arg Gly Phe Arg Arg Val Leu Asn Gln Arg Val Asp Ala Tyr  
15 20 25

TTT	GCC	GAG	CAT	GGC	CTG	ACC	CAA	AGG	GAT	AAT	CCC	TCC	ATG	TAT	CTG	2127
Phe	Ala	Glu	His	Gly	Leu	Thr	Gln	Arg	Asp	Asn	Pro	Ser	Met	Tyr	Leu	
			30					35					40			
AAA	ACC	CTG	ATT	ATT	GTG	CTC	TGG	TTG	TTT	TCC	GCT	TGG	GCC	TTT	GTG	2175
Lys	Thr	Leu	Ile	Ile	Val	Leu	Trp	Leu	Phe	Ser	Ala	Trp	Ala	Phe	Val	
		45					50					55				
CTT	TTT	GCT	CCA	GTT	ATT	TTT	CCG	GTG	CGC	CTA	CTG	GGT	TGT	ATG	GTT	2223
Leu	Phe	Ala	Pro	Val	Ile	Phe	Pro	Val	Arg	Leu	Leu	Gly	Cys	Met	Val	
	60					65					70					
TTG	GCG	ATC	GCC	TTG	GCG	GCC	TTT	TCC	TTC	AAT	GTC	GGC	CAC	GAT	GCC	2271
Leu	Ala	Ile	Ala	Leu	Ala	Ala	Phe	Ser	Phe	Asn	Val	Gly	His	Asp	Ala	
75					80					85					90	
AAC	CAC	AAT	GCC	TAT	TCC	TCC	AAT	CCC	CAC	ATC	AAC	CGG	GTT	CTG	GGC	2319
Asn	His	Asn	Ala	Tyr	Ser	Ser	Asn	Pro	His	Ile	Asn	Arg	Val	Leu	Gly	
				95					100					105		
ATG	ACC	TAC	GAT	TTT	GTC	GGG	TTA	TCT	AGT	TTT	CTT	TGG	CGC	TAT	CGC	2367
Met	Thr	Tyr	Asp	Phe	Val	Gly	Leu	Ser	Ser	Phe	Leu	Trp	Arg	Tyr	Arg	
			110					115					120			
CAC	AAC	TAT	TTG	CAC	CAC	ACC	TAC	ACC	AAT	ATT	CTT	GGC	CAT	GAC	GTG	2415
His	Asn	Tyr	Leu	His	His	Thr	Tyr	Thr	Asn	Ile	Leu	Gly	His	Asp	Val	
		125				130						135				
GAA	ATC	CAT	GGA	GAT	GGC	GCA	GTA	CGT	ATG	AGT	CCT	GAA	CAA	GAA	CAT	2463
Glu	Ile	His	Gly	Asp	Gly	Ala	Val	Arg	Met	Ser	Pro	Glu	Gln	Glu	His	
140						145					150					
GTT	GGT	ATT	TAT	CGT	TTC	CAG	CAA	TTT	TAT	ATT	TGG	GGT	TTA	TAT	CTT	2511
Val	Gly	Ile	Tyr	Arg	Phe	Gln	Gln	Phe	Tyr	Ile	Trp	Gly	Leu	Tyr	Leu	
155					160					165					170	
TTC	ATT	CCC	TTT	TAT	TGG	TTT	CTC	TAC	GAT	GTC	TAC	CTA	GTG	CTT	AAT	2559
Phe	Ile	Pro	Phe	Tyr	Trp	Phe	Leu	Tyr	Asp	Val	Tyr	Leu	Val	Leu	Asn	
				175					180					185		
AAA	GGC	AAA	TAT	CAC	GAC	CAT	AAA	ATT	CCT	CCT	TTC	CAG	CCC	CTA	GAA	2607
Lys	Gly	Lys	Tyr	His	Asp	His	Lys	Ile	Pro	Pro	Phe	Gln	Pro	Leu	Glu	
			190					195					200			
TTA	GCT	AGT	TTG	CTA	GGG	ATT	AAG	CTA	TTA	TGG	CTC	GGC	TAC	GTT	TTC	2655
Leu	Ala	Ser	Leu	Leu	Gly	Ile	Lys	Leu	Leu	Trp	Leu	Gly	Tyr	Val	Phe	
		205					210					215				
GGC	TTA	CCT	CTG	GCT	CTG	GGC	TTT	TCC	ATT	CCT	GAA	GTA	TTA	ATT	GGT	2703
Gly	Leu	Pro	Leu	Ala	Leu	Gly	Phe	Ser	Ile	Pro	Glu	Val	Leu	Ile	Gly	
220						225					230					
GCT	TCG	GTA	ACC	TAT	ATG	ACC	TAT	GGC	ATC	GTG	GTT	TGC	ACC	ATC	TTT	2751
Ala	Ser	Val	Thr	Tyr	Met	Thr	Tyr	Gly	Ile	Val	Val	Cys	Thr	Ile	Phe	
235					240					245					250	

ATG CTG GCC CAT GTG TTG GAA TCA ACT GAA TTT CTC ACC CCC GAT GGT	2799
Met Leu Ala His Val Leu Glu Ser Thr Glu Phe Leu Thr Pro Asp Gly	
255 260 265	
GAA TCC GGT GCC ATT GAT GAC GAG TGG GCT ATT TGC CAA ATT CGT ACC	2847
Glu Ser Gly Ala Ile Asp Asp Glu Trp Ala Ile Cys Gln Ile Arg Thr	
270 275 280	
ACG GCC AAT TTT GCC ACC AAT AAT CCC TTT TGG AAC TGG TTT TGT GGC	2895
Thr Ala Asn Phe Ala Thr Asn Asn Pro Phe Trp Asn Trp Phe Cys Gly	
285 290 295	
GGT TTA AAT CAC CAA GTT ACC CAC CAT CTT TTC CCC AAT ATT TGT CAT	2943
Gly Leu Asn His Gln Val Thr His His Leu Phe Pro Asn Ile Cys His	
300 305 310	
ATT CAC TAT CCC CAA TTG GAA AAT ATT ATT AAG GAT GTT TGC CAA GAG	2991
Ile His Tyr Pro Gln Leu Glu Asn Ile Ile Lys Asp Val Cys Gln Glu	
315 320 325 330	
TTT GGT GTG GAA TAT AAA GTT TAT CCC ACC TTC AAA GCG GCG ATC GCC	3039
Phe Gly Val Glu Tyr Lys Val Tyr Pro Thr Phe Lys Ala Ala Ile Ala	
335 340 345	
TGT AAC TAT CGC TGG CTA GAG GCC ATG GGC AAA GCA TCG TGACATTGCC	3088
Ser Asn Tyr Arg Trp Leu Glu Ala Met Gly Lys Ala Ser	
350 355 360	
TTGGGATTGA AGCAAAATGG CAAAATCCCT CGTAAATCTA TGATCGAAGC CTTTCTGTTG	3148
CCCGCCGACC AAATCCCCGA TGCTGACCAA AGGTTGATGT TGGCATTGCT CCAAACCCAC	3208
TTTGAGGGGG TTCATTGGCC GCAGTTTCAA GCTGACCTAG GAGGCAAAGA TTGGGTGATT	3268
TTGCTCAAAT CCGCTGGGAT ATTGAAAGGC TTCACCACCT TTGGTTTCTA CCCTGCTCAA	3328
TGGAAGGAC AAACCGTCAG AATTGTTTAT TCTGGTGACA CCATCACCGA CCCATCCATG	3388
TGGTCTAACC CAGCCCTGGC CAAGGCTTGG ACCAAGGCCA TGCAAATTCT CCACGAGGCT	3448
AGGCCAGAAA AATTATATTG GCTCCTGATT TCTTCCGGCT ATCGCACCTA CCGATTTTGT	3508
AGCATTTTGT CCAAGGAATT CTATCCCCAC TATCTCCATC CCACTCCCCC GCCTGTACAA	3568
AATTTTATCC ATCAGCTAGC	3588

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 359 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Thr Ala Glu Arg Ile Lys Phe Thr Gln Lys Arg Gly Phe Arg  
1 5 10 15  
Arg Val Leu Asn Gln Arg Val Asp Ala Tyr Phe Ala Glu His Gly Leu  
20 25 30  
Thr Gln Arg Asp Asn Pro Ser Met Tyr Leu Lys Thr Leu Ile Ile Val  
35 40 45  
Leu Trp Leu Phe Ser Ala Trp Ala Phe Val Leu Phe Ala Pro Val Ile  
50 55 60  
Phe Pro Val Arg Leu Leu Gly Cys Met Val Leu Ala Ile Ala Leu Ala  
65 70 75 80  
Ala Phe Ser Phe Asn Val Gly His Asp Ala Asn His Asn Ala Tyr Ser  
85 90 95  
Ser Asn Pro His Ile Asn Arg Val Leu Gly Met Thr Tyr Asp Phe Val  
100 105 110  
Gly Leu Ser Ser Phe Leu Trp Arg Tyr Arg His Asn Tyr Leu His His  
115 120 125  
Thr Tyr Thr Asn Ile Leu Gly His Asp Val Glu Ile His Gly Asp Gly  
130 135 140  
Ala Val Arg Met Ser Pro Glu Gln Glu His Val Gly Ile Tyr Arg Phe  
145 150 155 160  
Gln Gln Phe Tyr Ile Trp Gly Leu Tyr Leu Phe Ile Pro Phe Tyr Trp  
165 170 175  
Phe Leu Tyr Asp Val Tyr Leu Val Leu Asn Lys Gly Lys Tyr His Asp  
180 185 190  
His Lys Ile Pro Pro Phe Gln Pro Leu Glu Leu Ala Ser Leu Leu Gly  
195 200 205  
Ile Lys Leu Leu Trp Leu Gly Tyr Val Phe Gly Leu Pro Leu Ala Leu  
210 215 220  
Gly Phe Ser Ile Pro Glu Val Leu Ile Gly Ala Ser Val Thr Tyr Met  
225 230 235 240  
Thr Tyr Gly Ile Val Val Cys Thr Ile Phe Met Leu Ala His Val Leu  
245 250 255  
Glu Ser Thr Glu Phe Leu Thr Pro Asp Gly Glu Ser Gly Ala Ile Asp  
260 265 270

Asp Glu Trp Ala Ile Cys Gln Ile Arg Thr Thr Ala Asn Phe Ala Thr  
275 280 285

Asn Asn Pro Phe Trp Asn Trp Phe Cys Gly Gly Leu Asn His Gln Val  
290 295 300

Thr His His Leu Phe Pro Asn Ile Cys His Ile His Tyr Pro Gln Leu  
305 310 315 320

Glu Asn Ile Ile Lys Asp Val Cys Gln Glu Phe Gly Val Glu Tyr Lys  
325 330 335

Val Tyr Pro Thr Phe Lys Ala Ala Ile Ala Ser Asn Tyr Arg Trp Leu  
340 345 350

Glu Ala Met Gly Lys Ala Ser  
355

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1884 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGGTTCACTT CGGTTTTATA TTGTGACCAT GGTTCACAGG CATCTGCTCT AGGGAGTTTT	60
TCCGCTGCCT TTAGAGAGTA TTTTCTCCAA GTCGGCTAAC TCCCCCATTT TTAGGCAAAA	120
TCATATACAG ACTATCCCAA TATTGCCAGA GCTTTGATGA CTCACTGTAG AAGGCAGACT	180
AAAATTCTAG CAATGGACTC CCAGTTGGAA TAAATTTTAA GTCTCCCCCG GCGCTGGAGT	240
TTTTTTGTAG TTAATGGCGG TATAATGTGA AAGTTTTTTA TCTATTTAAA TTTATAAATG	300
CTAACAGCGG AAAGAATTAA ATTTACCCAG AACGGGGGT TTCGTCGGGT ACTAAACCAA	360
CGGGTGGATG CCTACTTTGC CGAGCATGGC CTGACCCAAA GGGATAATCC CTCCATGTAT	420
CTGAAAACCC TGATTATTGT GCTCTGGTTG TTTTCCGCTT GGGCCTTTGT GCTTTTTGCT	480
CCAGTTATTT TTCCGGTGCG CCTACTGGGT TGTATGGTTT TGGCGATCGC CTTGGCGGCC	540
TTTTCTTCA ATGTCGGCCA CGATGCCAAC CACAATGCCT ATTCCTCCAA TCCCCACATC	600
AACCGGGTTC TGGGCATGAC CTACGATTTT GTCGGGTAT CTAGTTTTCT TTGGCGCTAT	660
CGCCACAACCT ATTTGCACCA CACCTACACC AATATTCTTG GCCATGACGT GGAAATCCAT	720

GGAGATGGCG CAGTACGTAT GAGTCCTGAA CAAGAACATG TTGGTATTTA TCGTTTCCAG	780
CAATTTTATA TTTGGGGTTT ATATCTTTTC ATTCCCTTTT ATTGGTTTCT CTACGATGTC	840
TACCTAGTGC TTAATAAAGG CAAATATCAC GACCATAAAA TTCCTCCTTT CCAGCCCCTA	900
GAATTAGCTA GTTTGCTAGG GATTAAGCTA TTATGGCTCG GCTACGTTTT CGGCTTACCT	960
CTGGCTCTGG GCTTTTCCAT TCCTGAAGTA TTAATTGGTG CTTCGGTAAC CTATATGACC	1020
TATGGCATCG TGGTTTGCAC CATCTTTATG CTGGCCCATG TGTTGGAATC AACTGAATTT	1080
CTCACCCCCG ATGGTGAATC CGGTGCCATT GATGACGAGT GGGCTATTTG CCAAATTCGT	1140
ACCACGGCCA ATTTTGCCAC CAATAATCCC TTTTGGAACT GGTTTTGTGG CGGTTTAAAT	1200
CACCAAGTTA CCCACCATCT TTTCCCCAAT ATTTGTCATA TTCACTATCC CCAATTGGAA	1260
AATATTATTA AGGATGTTTG CCAAGAGTTT GGTGTGGAAT ATAAAGTTTA TCCCACCTTC	1320
AAAGCGGCGA TCGCCTCTAA CTATCGCTGG CTAGAGGCCA TGGGCAAAGC ATCGTGACAT	1380
TGCCTTGGA TTGAAGCAAA ATGGCAAAAT CCCTCGTAAA TCTATGATCG AAGCCTTTCT	1440
GTGCCCCGCC GACCAAATCC CCGATGCTGA CCAAAGGTTG ATGTTGGCAT TGCTCCAAAC	1500
CCACTTTGAG GGGGTTTCATT GGCCGCAGTT TCAAGCTGAC CTAGGAGGCA AAGATTGGGT	1560
GATTTTGCTC AAATCCGCTG GGATATTGAA AGGCTTCACC ACCTTTGGTT TCTACCCTGC	1620
TCAATGGGAA GGACAAACCG TCAGAATTGT TTATTCTGGT GACACCATCA CCGACCCATC	1680
CATGTGGTCT AACCAGCCC TGGCCAAGGC TTGGACCAAG GCCATGCAAA TTCTCCACGA	1740
GGCTAGGCCA GAAAAATTAT ATTGGCTCCT GATTTCTTCC GGCTATCGCA CCTACCGATT	1800
TTTGAGCATT TTTGCCAAGG AATTCTATCC CCACTATCTC CATCCCCTC CCCC GCCTGT	1860
ACAAAATTTT ATCCATCAGC TAGC	1884

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1685 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATATCTGCC TACCCTCCCA AAGAGAGTAG TCATTTTTC TCAATGGCTG CTCAAATCAA	60
--	----



GAAATACATT ACCTCAGATG AACTCAAGAA CCACGATAAA CCCGGAGATC TATGGATCTC	120
GATTCAAGGG AAAGCCTATG ATGTTTCGGA TTGGGTGAAA GACCATCCAG GTGGCAGCTT	180
TCCCTTGAAG AGTCTTGCTG GTCAAGAGGT AACTGATGCA TTTGTTGCAT TCCATCCTGC	240
CTCTACATGG AAGAATCTTG ATAAGTTTTT CACTGGGTAT TATCTTAAAG ATTACTCTGT	300
TTCTGAGGTT TCTAAAGATT ATAGGAAGCT TGTGTTTGAG TTTTCTAAAA TGGGTTTGTA	360
TGACAAAAAA GGTCATATTA TGTTTGCAAC TTTGTGCTTT ATAGCAATGC TGTTTGCTAT	420
GAGTGTTTAT GGGGTTTTGT TTTGTGAGGG TGTTTTGGTA CATTTGTTTT CTGGGTGTTT	480
GATGGGGTTT CTTTGGATTC AGAGTGGTTG GATTGGACAT GATGCTGGGC ATTATATGGT	540
AGTGTCTGAT TCAAGGCTTA ATAAGTTTAT GGGTATTTTT GCTGCAAATT GTCTTTCAGG	600
AATAAGTATT GGTTGGTGGA AATGGAACCA TAATGCACAT CACATTGCCT GTAATAGCCT	660
TGAATATGAC CCTGATTTAC AATATATACC ATTCCTTGTT GTGTCTTCCA AGTTTTTTGG	720
TTCACTCACC TCTCATTTCT ATGAGAAAAG GTTGACTTTT GACTCTTTAT CAAGATTCTT	780
TGTAAGTTAT CAACATTGGA CATTTTACCC TATTATGTGT GCTGCTAGGC TCAATATGTA	840
TGTACAATCT CTCATAATGT TGTTGACCAA GAGAAATGTG TCCTATCGAG CTCAGGAACT	900
CTTGGGATGC CTAGTGTTCT CGATTTGGTA CCCGTTGCTT GTTTCTTGTT TGCCTAATTG	960
GGGTGAAAGA ATTATGTTTG TTATTGCAAG TTTATCAGTG ACTGGAATGC AACAAAGTTCA	1020
GTTCTCCTTG AACCACTTCT CTTCAAGTGT TTATGTTGGA AAGCCTAAAG GGAATAATTG	1080
GTTTGAGAAA CAAACGGATG GGACACTTGA CATTTCTTGT CCTCCTTGGA TGGATTGGTT	1140
TCATGGTGGA TTGCAATTCC AAATTGAGCA TCATTTGTTT CCCAAGATGC CTAGATGCAA	1200
CCTTAGGAAA ATCTCGCCCT ACGTGATCGA GTTATGCAAG AAACATAATT TGCCTTACAA	1260
TTATGCATCT TTCTCCAAGG CCAATGAAAT GACACTCAGA ACATTGAGGA ACACAGCATT	1320
GCAGGCTAGG GATATAACCA AGCCGCTCCC GAAGAATTG GTATGGGAAG CTCTTCACAC	1380
TCATGGTTAA AATTACCCTT AGTTCATGTA ATAATTTGAG ATTATGTATC TCCTATGTTT	1440
GTGTCTTGTC TTGGTTCTAC TTGTTGGAGT CATTGCAACT TGTCTTTTAT GGTTTATTAG	1500
ATGTTTTTTA ATATATTTTA GAGGTTTTGC TTTCATCTCC ATTATTGATG AATAAGGAGT	1560
TGCATATTGT CAATTGTTGT GCTCAATATC TGATATTTTG GAATGTACTT TGTACCACTG	1620
TGTTTTCAGT TGAAGCTCAT GTGTA CTCT ATAGACTTTG TTAAATGGT TATGTCATGT	1680
TATTT	1685

789	
790	100
791	100
792	100
793	100
794	100
795	100
796	100
797	100
798	100
799	100
800	100
801	100
802	100
803	100
804	100
805	100
806	100
807	100
808	100
809	100
810	100
811	100
812	100
813	100
814	100
815	100
816	100
817	100
818	100
819	100
820	100
821	100
822	100
823	100
824	100
825	100
826	100
827	100
828	100
829	100
830	100
831	100
832	100
833	100
834	100
835	100
836	100
837	100
838	100
839	100
840	100
841	100
842	100
843	100
844	100
845	100
846	100
847	100
848	100
849	100
850	100
851	100
852	100
853	100
854	100
855	100
856	100
857	100
858	100
859	100
860	100
861	100
862	100
863	100
864	100
865	100
866	100
867	100
868	100
869	100
870	100
871	100
872	100
873	100
874	100
875	100
876	100
877	100
878	100
879	100
880	100
881	100
882	100
883	100
884	100
885	100
886	100
887	100
888	100
889	100
890	100
891	100
892	100
893	100
894	100
895	100
896	100
897	100
898	100
899	100
900	100

10029755-122101

Phe	Gly	Ser	Leu	Thr	Ser	His	Phe	Tyr	Glu	Lys	Arg	Leu	Thr	Phe	Asp	225	230	235	240
Ser	Leu	Ser	Arg	Phe	Phe	Val	Ser	Tyr	Gln	His	Trp	Thr	Phe	Tyr	Pro	245	250	255	
Ile	Met	Cys	Ala	Ala	Arg	Leu	Asn	Met	Tyr	Val	Gln	Ser	Leu	Ile	Met	260	265	270	
Leu	Leu	Thr	Lys	Arg	Asn	Val	Ser	Tyr	Arg	Ala	Gln	Glu	Leu	Leu	Gly	275	280	285	
Cys	Leu	Val	Phe	Ser	Ile	Trp	Tyr	Pro	Leu	Leu	Val	Ser	Cys	Leu	Pro	290	295	300	
Asn	Trp	Gly	Glu	Arg	Ile	Met	Phe	Val	Ile	Ala	Ser	Leu	Ser	Val	Thr	305	310	315	320
Gly	Met	Gln	Gln	Val	Gln	Phe	Ser	Leu	Asn	His	Phe	Ser	Ser	Ser	Val	325	330	335	
Tyr	Val	Gly	Lys	Pro	Lys	Gly	Asn	Asn	Trp	Phe	Glu	Lys	Gln	Thr	Asp	340	345	350	
Gly	Thr	Leu	Asp	Ile	Ser	Cys	Pro	Pro	Trp	Met	Asp	Trp	Phe	His	Gly	355	360	365	
Gly	Ser	Gln	Phe	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Lys	Met	Pro	Arg	370	375	380	
Cys	Asn	Leu	Arg	Lys	Ile	Ser	Pro	Tyr	Val	Ile	Glu	Leu	Cys	Lys	Lys	385	390	395	400
His	Asn	Leu	Pro	Tyr	Asn	Tyr	Ala	Ser	Phe	Ser	Lys	Ala	Asn	Glu	Met	405	410	415	
Thr	Leu	Arg	Thr	Leu	Arg	Asn	Thr	Ala	Leu	Gln	Ala	Arg	Asp	Ile	Thr	420	425	430	
Lys	Pro	Leu	Pro	Lys	Asn	Leu	Val	Trp	Glu	Ala	Leu	His	Thr	His	Gly	435	440	445	

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Trp Ile Gly His Asp Ala Gly His  
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asn Val Gly His Asp Ala Asn His  
1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Leu Gly His Asp Cys Gly His  
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (peptide)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Val Ile Ala His Glu Cys Gly His  
1 5

10029756-122101  
TOT221 95/62004

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Ile Gly His Asp Cys Ala His  
1 5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Val Val Gly His Asp Cys Gly His  
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

His Asn Ala His His  
1 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (peptide)

10029756 123101  
101221 95262001

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

His Asn Tyr Leu His His  
1 5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His Arg Thr His His  
1 5

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

His Arg Arg His His  
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

His Asp Arg His His  
1 5

10029756-122101

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

His Asp Gln His His  
1 5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

His Asp His His His  
1 5

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

His Asn His His His  
1 5

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

10029756.122101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Phe Gln Ile Glu His His  
1 5

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

His Gln Val Thr His His  
1 5

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

His Val Ile His His  
1 5

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

His Val Ala His His  
1 5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids

10029756-122101



(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

His Ile Pro His His  
1 5

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

His Val Pro His His  
1 5

2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1702 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 48..1406

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 48..1406

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CCCCAAAAT TTTCATTGTT CTCCATCTGG ACCACAGCAT CCACACA ATG GAG GGC	56
Met Glu Gly	
1	
GAA GCT AAG AAG TAT ATC ACG GCG GAG GAC CTC CGC CGC CAC AAC AAG	104
Glu Ala Lys Lys Tyr Ile Thr Ala Glu Asp Leu Arg Arg His Asn Lys	
5 10 15	
TCC GGC GAT CTC TGG ATC TCC ATC CAG GGC AAG GTC TAC GAC TGC TCT	152

Ser 20	Gly	Asp	Leu	Trp	Ile 25	Ser	Ile	Gln	Gly	Lys 30	Val	Tyr	Asp	Cys	Ser 35	
CGG	TGG	GCG	GCG	GAG	CAC	CCC	GGC	GGC	GAG	GTC	CCG	CTC	CTC	AGT	CTG	200
Arg	Trp	Ala	Ala	Glu 40	His	Pro	Gly	Gly	Glu 45	Val	Pro	Leu	Leu	Ser 50	Leu	
GCC	GGC	CAG	GAC	GTC	ACC	GAC	GCC	TTC	ATT	GCG	TAC	CAC	CCG	GGC	ACG	248
Ala	Gly	Gln	Asp 55	Val	Thr	Asp	Ala	Phe 60	Ile	Ala	Tyr	His	Pro 65	Gly	Thr	
GCG	TGG	CGG	CAT	CTG	GAT	CCG	CTC	TTC	ACC	GGC	TAC	TAC	TAC	CTC	AAG	296
Ala	Trp	Arg 70	His	Leu	Asp	Pro	Leu 75	Phe	Thr	Gly	Tyr	Tyr 80	Tyr	Leu	Lys	
GAC	TTC	GAA	GTG	TCG	GAG	ATC	TCC	AAG	GAC	TAC	CGG	AGG	CTT	TTG	AAC	344
Asp	Phe 85	Glu	Val	Ser	Glu	Ile 90	Ser	Lys	Asp	Tyr	Arg 95	Arg	Leu	Leu	Asn	
GAG	ATG	TCG	CGG	TCC	GGG	ATC	TTC	GAG	AAG	AAG	GGC	CAC	CAC	ATC	ATG	392
Glu	Met	Ser	Arg	Ser	Gly 105	Ile	Phe	Glu	Lys 110	Lys	Gly	His	His	Ile 115	Met	
TGG	ACG	TTC	GTC	GGC	GTT	GCG	GTC	ATG	ATG	GCG	GCA	ATC	GTC	TAC	GGC	440
Trp	Thr	Phe	Val	Gly 120	Val	Ala	Val	Met	Met 125	Ala	Ala	Ile	Val	Tyr 130	Gly	
GTG	CTG	GCG	TCG	GAG	TCC	GTC	GGA	GTT	CAC	ATG	CTC	TGC	GGC	GCA	CTG	488
Val	Leu	Ala	Ser 135	Glu	Ser	Val	Gly	Val 140	His	Met	Leu	Cys	Gly 145	Ala	Leu	
CTG	GGC	TTG	CTG	TGG	ATC	CAA	GCC	GCG	TAT	GTG	GGC	CAT	GAC	TCC	GGC	536
Leu	Gly	Leu	Leu	Trp	Ile	Gln	Ala 155	Ala	Tyr	Val	Gly	His 160	Asp	Ser	Gly	
CAT	TAC	CAG	GTG	ATG	CCA	ACC	CGT	GGA	TAC	AAC	AGA	ATC	ACG	CAA	CTC	584
His	Tyr 165	Gln	Val	Met	Pro	Thr 170	Arg	Gly	Tyr	Asn 175	Arg	Ile	Thr	Gln	Leu	
ATA	GCA	GGC	AAC	ATC	CTA	ACC	GGA	ATC	AGC	ATC	GCG	TGG	TGG	AAG	TGG	632
Ile	Ala	Gly	Asn	Ile	Leu 185	Thr	Gly	Ile	Ser	Ile 190	Ala	Trp	Trp	Lys	Trp 195	
ACC	CAC	AAC	GCC	CAC	CAC	CTC	GCC	TGC	AAC	AGC	CTC	GAC	TAC	GAC	CCC	680
Thr	His	Asn	Ala	His 200	His	Leu	Ala	Cys	Asn 205	Ser	Leu	Asp	Tyr	Asp 210	Pro	
GAC	CTC	CAG	CAC	ATC	CCC	GTA	TTC	GCC	GTC	TCC	ACC	CGA	CTC	TTC	AAC	728
Asp	Leu	Gln	His 215	Ile	Pro	Val	Phe	Ala 220	Val	Ser	Thr	Arg	Leu 225	Phe	Asn	
TCC	ATC	ACC	TCG	GTC	TTC	TAT	GGC	CGA	GTC	CTG	AAA	TTC	GAC	GAA	GTG	776
Ser	Ile	Thr 230	Ser	Val	Phe	Tyr	Gly 235	Arg	Val	Leu	Lys	Phe 240	Asp	Glu	Val	

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

CCACAATATT	GAAGTGAATA	ACCATGGAAG	GCACTACGTT	CAGCTTAACT	TTGCTTAACT	1516
TTGCTAGCTG	GTTGCGTTCC	CTTGTTGGGG	GCAAAGTGCA	GTATTTATTT	TCTTATCCCA	1576
TGTACTTTTT	GATTATTGTT	CTTATTCGTA	TCATAAATAA	TTTATTATTG	ATTAATTTTT	1636
GTTGTAGTTG	GGTGTCTATA	GCAAGTTTAT	AATACTGAGA	TATATTTTTT	TGGTAAAAAA	1696
AAAAAA						1702

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met 1	Glu	Gly	Glu	Ala 5	Lys	Lys	Tyr	Ile	Thr 10	Ala	Glu	Asp	Leu	Arg 15	Arg	
His	Asn	Lys	Ser 20	Gly	Asp	Leu	Trp	Ile 25	Ser	Ile	Gln	Gly	Lys 30	Val	Tyr	
Asp	Cys 35	Ser	Arg	Trp	Ala	Ala	Glu 40	His	Pro	Gly	Gly	Glu 45	Val	Pro	Leu	
Leu	Ser 50	Leu	Ala	Gly	Gln	Asp 55	Val	Thr	Asp	Ala	Phe 60	Ile	Ala	Tyr	His	
Pro 65	Gly	Thr	Ala	Trp	Arg 70	His	Leu	Asp	Pro	Leu 75	Phe	Thr	Gly	Tyr	Tyr 80	
Tyr	Leu	Lys	Asp	Phe 85	Glu	Val	Ser	Glu	Ile 90	Ser	Lys	Asp	Tyr	Arg 95	Arg	
Leu	Leu	Asn	Glu 100	Met	Ser	Arg	Ser	Gly 105	Ile	Phe	Glu	Lys	Lys 110	Gly	His	
His	Ile	Met 115	Trp	Thr	Phe	Val	Gly 120	Val	Ala	Val	Met	Met 125	Ala	Ala	Ile	
Val	Tyr 130	Gly	Val	Leu	Ala	Ser 135	Glu	Ser	Val	Gly	Val 140	His	Met	Leu	Cys	
Gly 145	Ala	Leu	Leu	Gly	Leu 150	Leu	Trp	Ile	Gln	Ala 155	Ala	Tyr	Val	Gly	His 160	
Asp	Ser	Gly	His	Tyr 165	Gln	Val	Met	Pro	Thr 170	Arg	Gly	Tyr	Asn	Arg 175	Ile	

Thr Gln Leu Ile Ala Gly Asn Ile Leu Thr Gly Ile Ser Ile Ala Trp  
180 185 190

Trp Lys Trp Thr His Asn Ala His His Leu Ala Cys Asn Ser Leu Asp  
195 200 205

Tyr Asp Pro Asp Leu Gln His Ile Pro Val Phe Ala Val Ser Thr Arg  
210 215 220

Leu Phe Asn Ser Ile Thr Ser Val Phe Tyr Gly Arg Val Leu Lys Phe  
225 230 235 240

Asp Glu Val Ala Arg Phe Leu Val Ser Tyr Gln His Trp Thr Tyr Tyr  
245 250 255

Pro Val Met Ile Phe Gly Arg Val Asn Leu Phe Ile Gln Thr Phe Leu  
260 265 270

Leu Leu Leu Thr Arg Arg Asp Val Pro Asp Arg Ala Leu Asn Leu Met  
275 280 285

Gly Ile Ala Val Phe Trp Thr Trp Phe Pro Leu Phe Val Ser Cys Leu  
290 295 300

Pro Asn Trp Pro Glu Arg Phe Gly Phe Val Leu Ile Ser Phe Ala Val  
305 310 315 320

Thr Ala Ile Gln His Val Gln Phe Thr Leu Asn His Phe Ser Gly Asp  
325 330 335

Thr Tyr Val Gly Pro Pro Lys Gly Asp Asn Trp Phe Glu Lys Gln Thr  
340 345 350

Lys Gly Thr Ile Asp Ile Thr Cys Pro Pro Trp Met Asp Trp Phe Phe  
355 360 365

Gly Gly Leu Gln Phe Gln Leu Glu His His Leu Phe Pro Arg Leu Pro  
370 375 380

Arg Gly Gln Leu Arg Lys Ile Ala Pro Leu Ala Arg Asp Leu Cys Lys  
385 390 395 400

Lys His Gly Met Pro Tyr Arg Ser Phe Gly Phe Trp Asp Asp Ala Asn  
405 410 415

Val Arg Thr Ile Arg Thr Leu Arg Asp Ala Ala Val Gln Ala Arg Asp  
420 425 430

Leu Asn Ser Ala Pro Cys Pro Lys Lys Leu Gly Tyr Gly Glu Ala Tyr  
435 440 445

Asn Thr His Gly \*

450